

Epipolar geometry for vision-guided laser surgery.

Nicolas Andreff, Soukalo Dembélé, Brahim Tamadazte, Syed Zill-e Hussnain

Institut FEMTO-ST, Université de Franche-Comté/CNRS/ENSMM/UTBM, 24 rue Savary, 25000 Besançon, France

Nicolas.Andreff@femto-st.fr

Keywords: Laser Surgery ; Visual Servoing ; Epipolar geometry

Abstract: This paper proposes to use the analogy between a scanning laser beam and a camera. Thereby, a degenerate stereoscopic system can be defined by such a virtual camera and a real camera observing the laser spot on the tissues. Using the associated epipolar geometry, a control law is derived without any matrix inversion nor estimation of the 3D scene. This control law shows exponential convergence in simulation.

1 INTRODUCTION

The μ RALP project involves the development of a system for endoluminal laser phonosurgery, *i.e.* surgery of the vocal chords using a laser emitted from inside the larynx. Indeed, in current laryngeal laser surgical procedures, a beam of incision laser is projected on the target position of the soft tissue from the working distance of 400mm by means of a rigid laryngoscope (Fig.1). This yields safety concerns for the patient and staff, as well as limitations to accuracy. More, this so-called *laryngeal suspension* position of the patient requires an extreme extension of the neck, which makes it painful several days after the operation.

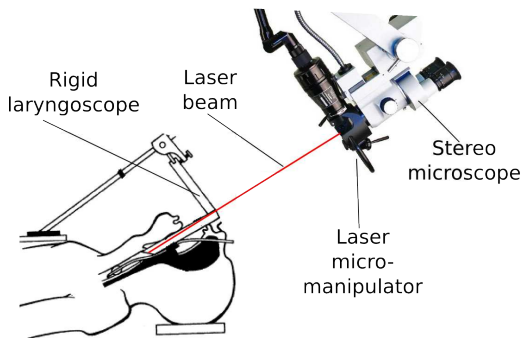


Figure 1: Current laryngeal laser surgery setup.

One such laryngeal laser surgical system is the AcuBladeTM. It claims to be a robotics system because it features fully autonomous pattern following (line for resection, circle for ablation) by the laser spot, once the surgeons has defined the dimension and position of the pattern intra-operatively based on direct visualization (through the microscope) of the op-

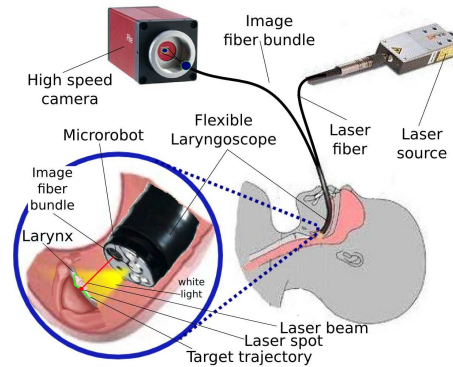


Figure 2: Endoscopic laser microphonosurgery

erating site. The laser spot is thus automatically controlled, which means that the objective of μ RALP to automate the control of the laser phonosurgery is totally in line with the industrial and medical practices in the field.

Yet, in the AcuBlade system, the laser spot is controlled in an open-loop mode from the surgeon's viewpoint. Indeed, once the surgeon has chosen the pattern parameters, the robot follows it blindly, relying on its internal calibration and sensors. Any deviation from this plan is forbidden (to the exception of the foot pedal switch of the power laser). However, during the ablation and even more during the resection, the tissues move and change. Therefore, the planned pattern must be very small and can not guarantee high accuracy. Moreover, in case of an endoscopic laser steering system (Fig. 2), one can not guarantee any time stability of the micro-robot calibration.

The only way to increase the accuracy and the extension of the patterns is to close the control loop over an exteroceptive sensor, namely the imaging system. Closing the loop over a camera is known as visual servoing and as being a technique robust to calibration errors and to changes in the environment of the robot.

Visual servoing has been used in minimally invasive surgery to perform two types of applications. The first type is relative to tissue motion tracking and compensation. By opposition to industrial objects human tissues are alive and than submitted to heart and breathing motions. These physiological motions complicate the task of the surgeon, and their compensation improve the precision of surgical interventions. The second type of applications deals with tracking and guidance of instruments. In this case, the objective is more classical for visual servoing: positioning of a target by using visual feedback.

1.1 Motion compensation

In motion compensation using image-based control, the motion of tissues is estimated from images of an imaging system and compensated by means of a robotized instrument.

Gangloff and his colleagues at Strasbourg have developed an active heart beating stabilizer by using a predictive control scheme (Ginhoux et al., 2005) and a robust control scheme (Bachta et al., 2011). They obtained interesting results with simulation experiments as well as in vivo experiments with pigs. In both cases the acquisition rate reached 500 Hz. In (Krupa et al., 2009) authors have used visual servoing and speckle information in ultrasound images at 12 Hz to stabilize a probe in a region of interest. Simulation experiments have been used to validate the developments. In (Kesner et al., 2010) authors have integrated heart motion compensation in the 3D positioning of a catheter using 3D ultrasound images. A Kalman filtering has allowed to take into account the 50-100 ms delay of acquisition and a tracking error of 0.77 mm was obtained with in vivo tests with pigs. In (Chen et al., 2010) a virtual bone clasper is achieved by means of visual servoing with a stereo camera running at 10 Hz and a Kalman filtering.

The core of these applications is motion tracking and for that solutions already developed in machine vision can be used: correlation (Ortmaier et al., 2005), Lucas-Kanade tracker or SURF (Elhawary and Popovic, 2011).

1.2 Instrument guidance

In instrument guidance by means of visual servoing many developments have been reported in the literature.

In 2003, Krupa et al. (Krupa et al., 2003) proposed the use of image-based visual servoing to perform 3D positioning of a surgical instrument. But only simulation experiments were performed with a camera running at 50 Hz and a phantom instrument. Krupa and his colleagues at Rennes have proposed the 3D positioning of a probe using visual servoing in ultrasound images: image-based visual servoing on a simulation setup in (Krupa and Chaumette, 2006), visual servoing using moments on simulation setup running at 25 Hz in (Mebarki et al., 2010).

In (Becker et al., 2010) the authors have developed a semiautomated laser intraocular surgery. Their setup included a laser probe attached to a micro-manipulator, a stereo camera working at 30 kHz, a 3D sensor running at 2 kHz and a control system running at 1 kHz. All the aspects of the application were presented: preoperative procedure for the selection of the sites to burn, calibration of stereo camera using the data of the 3D sensor, 3D surface reconstruction and registration of current images with preoperative images, PID control of the micro-manipulator for the positioning of the laser on the site to burn. This is a very interesting paper, which clearly can be useful for laser phonomicrosurgery.

In (Mattos and Caldwell, 2012) and (Dagnino et al., 2012) the team of Mattos at Genoa describe their experiments in laser phonomicrosurgery. They have developed a complete equipment including laser sources, micro-robots to control the position of the laser beam, an optical microscope viewing and recording images and a haptic device. They also realized algorithmic developments to follow the laser spot a predefined path. For the moment, they do treat the problem in a static point of view: they do not include the dynamics of the system (three-dimensional tissue, tissue movement, kinematic model of micro-robot, ...).

In (Reilink et al., 2010) the authors used directly images delivered by an endoscope to guide it inside the body. Simulation experiment show an improvement of 68% with respect to manual steering.

1.3 Related developments

In addition to above publications dealing with visual servoing directly, it is also necessary to cite the publications (Ota et al., 2009; Rivera-Serrano et al., 2012) which describe the development of a flexible robot for

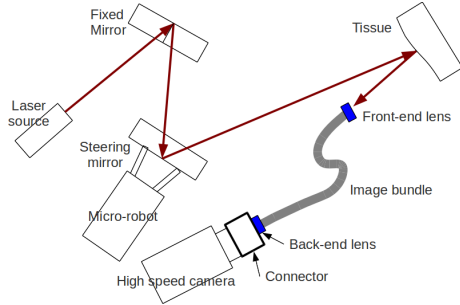


Figure 3: Schematic view of the laser steering system

transoral surgery. It is a snake-like mechanism 10 mm diameter, 300 mm long with 105 degree-of-freedom. It is endowed with a 15 K fiber bundle connected to a 640×480 pixels camera, an illumination lamp, two 4.2 mm tool ports. It has been possible to introduce it in cadavers without laryngeal suspension, to observe vocal cord and to perform a retraction and cauterization on the base of tongue. However, the system is only teleoperated which might limit its accuracy with respect to the delicate phonosurgery requirements. Such a system should thus be enhanced by adding visual servoing and, probably also, a laser for replacing the currently used mechanical scalpel.

1.4 Contribution of the paper

The contribution of this paper is to discuss the control of a laser (namely, the invisible incision laser for incision with co-axial visible HeNe laser pointer) over the vocal fold, and by extension, the control of a laser over any surface. It shows that making call to geometry can simplify the control: no matrix inversion, no explicit knowledge or reconstruction of the 3D scene.

2 VISUALLY-GUIDED LASER SURGERY

In this section, we analyse the control of the laser spot with the micro-robot using an optic fiber bundle to bring the image of the scene onto a high-speed camera (Fig. 2 and 3).

This control can be done in two ways: using the standard visual servoing equations or using the above grounding analogy.

2.1 A word on standard control

Let us note \underline{z} the direction of the laser beam reflecting from the steering mirror towards the vocal fold, P the

position of the laser spot on the tissue surface, and p the position of the laser spot in the image.

Then, it is trivial to write, in the reference frame R_0 attached to the zero-reference of the steering mirror:

$${}^0P = d {}^0\underline{z} \quad (1)$$

where d is the distance traveled by the laser from the mirror to the tissue. This distance can not be measure, to the contrary of ${}^0\underline{z}$ which can be obtained from the microrobot encoders. However, it can be modeled if one approximates the tissue surface in P by a plane of equation:

$${}^0\underline{n}^T {}^0P - d_0 = 0 \quad (2)$$

where ${}^0\underline{n}$ is the orientation of the surface normal in R_0 and d_0 is the distance of the plane to the origin of R_0 . Using this model, one finds:

$$d = \frac{d_0}{{}^0\underline{n}^T {}^0\underline{z}} \quad (3)$$

On the other hand, the perspective projection equation yields:

$$\tilde{p} = \mathbf{K} \frac{{}^cP}{Z} \quad (4)$$

where \mathbf{K} is the matrix containing the intrinsic parameters, P is now expressed in the camera frame R_c , Z is the unmeasured depth along the line of sight passing through p as well as the third coordinate of cP and \tilde{p} represents the homogeneous coordinates of p .

To apply, the usual visual servoing approach, one needs to differentiate the latter with time:

$$\dot{\tilde{p}} = \frac{1}{Z} \mathbf{K} (\mathbf{I}_2 - p) {}^c\dot{P} \quad (5)$$

One has another expression for ${}^c\dot{P}$ by differentiating (1):

$${}^0\dot{P} = d {}^0\dot{\underline{z}} + \dot{d} {}^0\underline{z} \quad (6)$$

and expressing the latter in R_c :

$${}^c\dot{P} = {}^c\mathbf{R}_0 (\dot{d} {}^0\underline{z} + d {}^0\dot{\underline{z}}) \quad (7)$$

Now, from 3, one gets (under the simplifying assumption that the surface plane does not change):

$$\dot{d} = - \frac{d_0 {}^0\underline{n}^T}{{}^0\underline{n}^T {}^0\underline{z}} {}^0\dot{\underline{z}} \quad (8)$$

Putting (7) and (8) into (5) gives therefore $\dot{\tilde{p}}$ under the form:

$$\dot{\tilde{p}} = \mathbf{L}(d, Z, {}^0\underline{n}, d_0, {}^c\mathbf{R}_0, p, {}^0\underline{z}) {}^0\dot{\underline{z}} \quad (9)$$

where \mathbf{L} is of dimension 3×3 . Inverting it allows to convert the image velocity of the laser spot into the velocity of the laser beam, which, in turn, shall be converted into microrobot velocity through the differential inverse kinematic model.

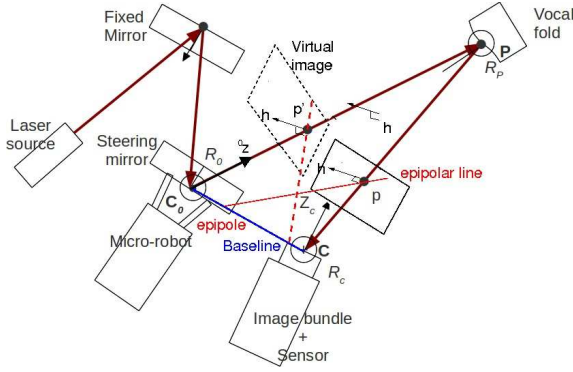


Figure 4: Analogy with stereoscopy.

Therefore, this controller needs the estimation of d and Z , which can be obtained by triangulation between the laser beam (known from the joint values q), the line of sight outgoing from the camera through the laser spot projection (p) and the robot-camera calibration.

We have not implemented this controller, nor pushed further the details of the calculation (so far), because there is a more elegant way of treating the problem, which does not require any explicit triangulation nor any matrix inversion.

2.2 Laser visual servoing using epipolar geometry

On the opposite to the above method, which is totally generic, the method we propose here is totally hand-made and tailored to the specific case of a laser beam being observed by a camera.

Indeed, the set-up in Fig. 3 is analogous to a degenerate case of epipolar geometry in Fig. 4. Thereby, points p and p' are the images of the same spatial point P and are hence linked by the epipolar constraint:

$$\tilde{p}'^\top \mathbf{F} \tilde{p} = 0 \quad (10)$$

where \mathbf{F} is the fundamental matrix of the 2-view system (Hartley and Zisserman, 2006). Actually, this epipolar constraint is defined up to a scale factor, and thereby, \tilde{p}' can be replaced by ${}^0\mathbf{z}$, the unit vector describing, in the microrobot baseframe R_0 , the direction of the laser beam from the mirror to the vocal fold:

$${}^0\mathbf{z}^\top \mathbf{F} \tilde{p} = 0 \quad (11)$$

This equation expresses the fact that the origin of the camera, the pivot point of the micro-robot, the laser beam, the line of sight and the laser spot on the tissue are coplanar. It can be also interpreted in three

ways:

$${}^0\mathbf{z} \perp \mathbf{F} \tilde{p} \quad (12)$$

$$\tilde{p} \perp \mathbf{F}^\top {}^0\mathbf{z} \quad (13)$$

and both $\mathbf{F} \tilde{p}$ and $\mathbf{F}^\top {}^0\mathbf{z}$ represent the (non-unit) normal vector to the epipolar plane, respectively in the micro-robot and the camera frame.

The time derivative of the epipolar constraint is:

$$(\mathbf{F}^\top {}^0\mathbf{z})^\top \dot{\tilde{p}} + (\mathbf{F} \tilde{p})^\top {}^0\dot{\mathbf{z}} = 0 \quad (14)$$

Now, we can decompose ${}^0\mathbf{z}$ into a component orthogonal to the epipolar plane and a component inside the latter:

$${}^0\mathbf{z} = \alpha {}^0\mathbf{h} + \beta {}^0\mathbf{z} \times {}^0\mathbf{h} \quad (15)$$

where ${}^0\mathbf{h} = \frac{\mathbf{F} \tilde{p}}{\|\mathbf{F} \tilde{p}\|}$. Replacing this expression into the epipolar constraint and reordering the terms, we get:

$$\alpha = -\frac{(\mathbf{F}^\top {}^0\mathbf{z})^\top}{\|\mathbf{F} \tilde{p}\|} \dot{\tilde{p}} \quad (16)$$

Actually, α only depends on the projection of $\dot{\tilde{p}}$ onto the normal to the epipolar plane, but expressed in the camera frame, *i.e.* along ${}^c\mathbf{h} = \frac{\mathbf{F}^\top {}^0\mathbf{z}}{\|\mathbf{F}^\top {}^0\mathbf{z}\|}$. Thus, the remaining part of $\dot{\tilde{p}}$ is obtained by cancelling this projection:

$$\dot{\tilde{p}} = a {}^c\mathbf{h} + (\mathbf{I}_3 - {}^c\mathbf{h} {}^c\mathbf{h}^\top) \dot{\tilde{p}} \quad (17)$$

where the value of a does not have any interest for the sequel, but can be related to α by inserting the latter equation into the former.

Now, concentrate on the part of $\dot{\tilde{p}}$ lying in the epipolar plane. In (17), it is expressed in the camera frame, so we just need to bring it back to the micro-robot frame, going backwards the camera intrinsic parameters, the orientation of the camera frame with respect to the micro-robot frame and compensating for the unknown scale factor in \mathbf{F} , to get β :

$$\beta = \frac{\|\mathbf{F}^\top {}^0\mathbf{z}\|}{\|\mathbf{F} \tilde{p}\|} ({}^0\mathbf{z} \times {}^0\mathbf{h})^\top {}^0\mathbf{R}_c \mathbf{K}^{-1} (\mathbf{I}_3 - {}^c\mathbf{h} {}^c\mathbf{h}^\top) \dot{\tilde{p}} \quad (18)$$

As a consequence, we have expressed ${}^0\dot{\mathbf{z}}$ as a function of $\dot{\tilde{p}}$:

$${}^0\dot{\mathbf{z}} = \left\{ ({}^0\mathbf{z} \times {}^0\mathbf{h}) ({}^0\mathbf{z} \times {}^0\mathbf{h})^\top {}^0\mathbf{R}_c \mathbf{K}^{-1} (\mathbf{I}_3 - {}^c\mathbf{h} {}^c\mathbf{h}^\top) - {}^0\mathbf{h} {}^c\mathbf{h}^\top \right\} \frac{\|\mathbf{F}^\top {}^0\mathbf{z}\|}{\|\mathbf{F} \tilde{p}\|} \dot{\tilde{p}} \quad (19)$$

Consequently, we have the exact expression of the conversion of the image velocity into the laser beam velocity, without any matrix inversion, nor any explicit triangulation or scene structure knowledge. This expression only depends on the measurements

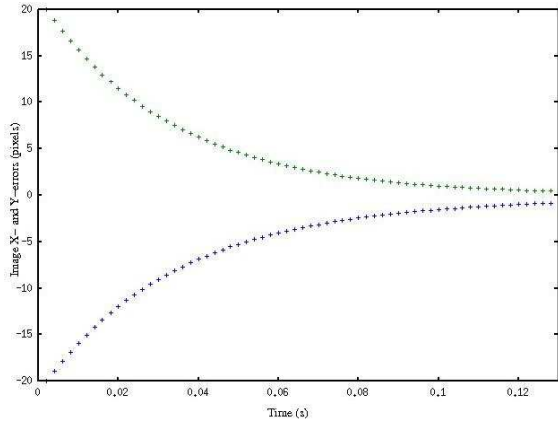


Figure 5: Time evolution of the image errors.

(${}^0\mathbf{z}$ and p), the fundamental matrix \mathbf{F} and a reduced set of calibration parameters (\mathbf{K} and ${}^c\mathbf{R}_0$). Geometry is always useful !

Now, we can come up to the control law, by enforcing a first order behaviour of the error in the image between the current and the desired projections of the laser spot:

$$\dot{\tilde{p}} = -\lambda(\tilde{p} - \tilde{p}^*) \quad (20)$$

3 SIMULATION RESULTS

This control was implemented in a simulator and gives exponential decay of the image error (Fig.??).

It shall be tested further with the introduction of noise in the sensor signals and the calibration parameters. Later, it shall be implemented on a basic hardware (standard camera + pan/tilt micropositionner from PI) before it is implemented in an endoscopic set-up.

4 DISCUSSION

In this paper, it was shown that resorting to geometry simplifies the eye-to-hand control law for a surgical laser (and any other application where a beam needs be accurately swept over a surface), by essentially removing the need for on-line estimation of the 3D surface. Actually, if a dedicated surgeon-robot interface is designed (?) to define the desired trajectory in the image, then the latter will geometrically contain a coherent description of the 3D surface, and thus, the 3D information is not purely and simply thrown away as it could seem but, rather, it is implicitly used.

Yet, the proposed proportional control in the image is essentially targeted at disturbance rejection pur-

pose. It should be complemented by an additional term for trajectory tracking purpose ($\tilde{p}^*(t)$), including constraints on the tissue exposure to laser, in order to avoid any carbonization.

Further investigation of geometry, namely the trifocal geometry associated to a stereoscopic observation of the laser spot, is expected to further simplify the control and to increase its robustness, which is a key issue in the transfer of automation into actual clinical devices.

Also, many micromanipulators have a parallel kinematics architecture, which are known to be controllable without any joint sensing (?). Consequently, using the proposed multi-view geometric approach might enable simplified miniaturization of laser steering in an endoscopic set-up, because one can design steering parallel kinematics mechanisms without proprioceptive sensors. This is a very crucial investigation field, since endoscopic laser surgery faces very contradictory requirements in terms of sweeping range and frequency (yielding larger mechanisms) and of available space at the endoscopic tip.

ACKNOWLEDGEMENTS

This work was supported by μ RALP, the EC FP7 ICT Collaborative Project no. 288663 (<http://www.microralp.eu>), and by ACTION, the French ANR Labex no. "ANR-11-LABX-01-01" (<http://www.labex-action.fr>).

REFERENCES

- Bachta, W., Renaud, P., Laroche, E., Forgione, A., and Gangloff, J. (2011). Active stabilization for robotized beating heart surgery. *IEEE Transactions on Robotics*, 27:757–568.
- Becker, B. C., MacLachlan, R. A., Jr, L. A. L., and Riviere, C. N. (2010). Semiautomated intraocular laser surgery using handheld instruments. *Lasers in Surgery and Medicine*, 42:264–273.
- Chen, C.-S., Hsieh, M.-S., Chiu, Y.-W., Tsai, C.-H., Liu, S.-M., Lu, C.-C., and Yen, P.-L. (2010). An unconstrained virtual bone clasper for knee surgical robot using visual servoing technique. *Journal of the Chinese Institute of Engineers*, 33(3):379–386.
- Dagnino, G., Mattos, L. S., and Caldwell, D. G. (2012). New software tools for enhanced precision in robot-assisted laser phonomicrosurgery. In *34th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC'12)*.
- Elhawary, H. and Popovic, A. (2011). Robust feature tracking on the beating heart for a robotic-guided endo-

- scope. *The international journal of medical robotics and computer assisted surgery*, 7:459–468.
- Ginhoux, R., Gangloff, J., de Mathelin, M., Soler, L., Sanchez, M. M. A., and Marescaux, J. (2005). Active filtering of physiological motion in robotized surgery using predictive control. *IEEE Transactions on Robotics and Automation*, 21(1):235–246.
- Hartley, R. and Zisserman, A. (2006). *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge, United Kingdom, 2nd edition edition.
- Kesner, S. B., Yuen, S. G., and Howe, R. D. (2010). Ultrasound servoing of catheters for beating heart valve repair. In *IPCAI*, pages 168–178.
- Krupa, A. and Chaumette, F. (2006). Guidance of an ultrasound probe by visual servoing. *Advanced Robotics*, 20(11):1203–1218.
- Krupa, A., Fichtinger, G., and Hager, G. D. (2009). Real-time motion stabilization with b-mode ultrasound using image speckle information and visual servoing. *International Journal of Robotics Research*.
- Krupa, A., Gangloff, J., Doignon, C., de Mathelin, M. F., Morel, G., Leroy, J., Soler, L., and Marescaux, J. (2003). Autonomous 3-d positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE Transaction on Robotics and Automation*, 19:842–853.
- Mattos, L. S. and Caldwell, D. G. (2012). Safe teleoperation based on flexible intraoperative planning for robot-assisted laser microsurgery. In *34th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC'12)*.
- Mebarki, R., Krupa, A., and Chaumette, F. (2010). 2-d ultrasound probe complete guidance by visual servoing using image moments. *IEEE Transactions on Robotics*, 26(2):296–306.
- Ortmaier, T., Groger, M., Boehm, D., Falk, V., and Hirzinger, G. (2005). Motion estimation in beating heart surgery. *IEEE Transactions on Biomedical Engineering*, 52(10):1729–1740.
- Ota, T., Degani, A., Schwartzman, D., Zubiate, B., McGarvey, J., Choset, H., and Zenati, M. A. (2009). A highly articulated robotic surgical system for minimally invasive surgery. *The Annals of Thoracic Surgery*, 87(4):1253–1256.
- Reilink, R., Stramigioli, S., and Misra, S. (2010). Image-based flexible endoscope steering. In *The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2339–2344.
- Rivera-Serrano, C. M., Johnson, P., Zubiate, B., Kuenzler, R., Choset, H., Zenati, M., Tully, S., and Duvvuri, U. (2012). A transoral highly flexible robot: Novel technology and application. *The Laryngoscope*, 122:1067–1071.